CHEMICAL AND MASS SPECTROMETRIC SEQUENCE STUDIES OF A PEPTIDE FROM THE VARIABLE PART OF NORMAL IMMUNOGLOBULIN λ-CHAINS *

F. FRANĚK and B. KEIL

Institute of Organic Chemistry and Biochemistry, Czechoslovak Academy of Sciences, Prague, Czechoslovakia

and

D. W. THOMAS and E. LEDERER

Institut de Chimie des Substances Naturelles, C.N.R.S., 91-Gif-sur-Yvette, France

Received 24 February 1969

1. Introduction

The polypeptide chains derived from normal immunoglobulins display a certain kind of microheterogeneity: in several positions more than one amino acid can be found. The determination of variable sequence meets with problems not fully considered upon studies of homogeneous proteins. Therefore a search for methods suitable to sequencing peptides possessing variable positions has a great significance. In this communication a comparative study is reported. The sequence of an octadecapeptide was independently determined on the one hand by sequential degradation (isothiocyanate method) combined with dansylation, and on the other by mass spectrometry. Up to ten residues from the *N*-terminus could be determined and the results of both these methods were in fair agreement.

2. Material and methods

The octadecapeptide T2bl was isolated from the tryptic hydrolysate of pig immunoglobulin λ-chains (ref. [1]).

Amino acid composition was determined according to Spackman and coworkers [2] on an amino acid analyzer Model 6020 (Developmental Workshops, Czechoslovak Academy of Sciences): 20-hr and 70-hr hydrolyzates were analyzed.

The isothiocyanate method for sequential degradation and the dansyl chloride procedure for determination of *N*-terminal amino acids was performed according to Gray [3,4]. Dansyl-amino acid derivatives were identified by thin-layer chromatography on silica gel [5].

The *N*-terminal acetyl derivative of the peptide was prepared with methanolic acetic anhydride as described earlier [6]. *O,N*-permethylation [7] of the derivative was performed with methyl iodide and sodium hydride in dimethylformamide as by Coggins and Benoiton [8].

Mass spectra were determined using the direct introduction probe of an A.E.I. model MS9 mass spectrometer.

3. Results and discussion

The amino acid composition of the octadecapeptide T2bl shows that some amino acids are present in amounts which differ significantly from integral numbers of residues, whereas a few amino acids were found

* Part XVII in the series "Determination of amino acid sequences in oligopeptides by mass spectrometry"; part XVI, D.W.Thomas and T.Ito, Tetrahedron, in press.

Table 1
Amino acid composition of peptide T2b1. (The calculation is based on the aminoethylcysteine content. Numbers of residues are given.)

Aminoethyl-cys	1.0	Val	0.1
Asp	2.2	Ile	1.0
Thr	2.4	Leu	1.1
Ser	0.1	Tyr	0.9
Glu	2.8	Phe	1.0
Gly	1.1		
Ala	4.4	Total	18.1

in quantities equimolar with the quantity of *C*-terminal aminoethylcysteine (table 1). The physical homogeneity of the peptide, as assayed by gel filtration, and the presence of a unique *N*-terminal amino acid (alanine) [1] indicate that abnormalities in amino acid composition are not necessarily due to admixtures of unrelated peptides, but more likely to amino acid replacements.

For sequential degradation, 2 mg of the peptide was used. With this amount the amino acids in positions 1 to 7 were determined. Another 1 mg served for determination of amino acids in positions 7 to 9 (no aliquot for dansylation was taken until 6th degradation step was completed). At least 50 nanomoles of peptide were taken for dansylation after each degradation step. Such an amount would be overmuch when handling a homogeneous peptide, but is necessary for the detection of minor amino acid replacements. Relative intensities of dansyl-amino acids found in the variable positions (table 2) can hardly be evaluated by inspection only. Therefore the vertical order of amino acids in positions 2, 4 and 6 has only an approximative character.

Table 2
Sequence of peptide T2b1 determined by sequential degradation.

Position:	1	2	3	4	5	6	7	8	9
Amino acid(s):	Ala	Thr	Leu	Thr	Ile	Thr	Gly	Ala	Glx
		Ala		Leu		Ala			
				Gly		Gly			

Where multiple amino acids were found in one position, the amino acid which seemed to be the most abundant is placed in the upper row.

To determine a sequence by mass spectrometry, it was first necessary to modify the octadecapeptide chemically in order to produce a derivative of greater volatility. For this purpose 5 mg of the peptide were acetylated in the *N*-terminal position. All hydroxyl groups of this derivative were left free. *O,N*-permethylation of this derivative gave a product which could be vaporized in the mass spectrometer at 350°. Electron-induced fragmentation of such a permethylated peptide derivative occurs principally at the Co-NMe bonds [9], and the charge remains on the fragment containing the *N*-terminus of the original peptide. The mass spectrum thus consists of a sequence of peaks represented by ions of the type:

$$\mathrm{Ac.NMeCHCO.....NMeCHCO.NMeC\overset{\oplus}{H}CO}$$
$$\qquad\qquad\quad R_1 \qquad\qquad\quad R_{n-1} \qquad\quad R_n$$

The major peaks in the spectrum of the peptide derivative at hand were found at the *m/e* values summarized in table 3. For the purpose of interpretation, one may consider the series of peaks of successively increasing *m/e* value to represent the stepwise "synthesis" of a peptide, starting at the *N*-terminus.

Thus, in table 3, the lowest *m/e* value of considerable intensity was *m/e* 128. Because the *N*-terminus was known to contain an acetyl function (mass 43), the mass of the first residue is 85 (128 minus 43), which corresponds uniquely to a residue of *N*-methylalanine. The next peak was found at *m/e* 213, and again an increment of 85 mass units (213 minus 128) demonstrated that the second residue was also a methylalanine.

The increment of 44 mass units between the second and third major peaks of this spectrum, *m/e* 213 and 257 respectively, does not correspond to any amino acid residue. Instead, the *m/e* 257 peak gives evidence for a mixture of residues at position 2: the increment of 129 mass units (257 minus 128) corresponds to a dimethylthreonine residue, and on the basis of peak intensities we could deduce that the peptide sample was an approximately equimolar mixture of at least two compounds (peptides A and B, table 3) which differ only in the amino acid composition at the second *N*-terminal residue.

A mass of 127 (340 minus 213 or 384 minus 257)

Table 3
Sequence determination of peptide T2bl by mass spectrometry.

Position	1	2	3	4	5	6	7	8	9	10
Peptide A:	Ac.MeAla.	MeAla.	MeLeu[a].	MeThr.	MeLeu[a].	MeThr.	MeGly.	MeAla.	MeGln.	MeAla.
				OMe		OMe			NMe_2	
m/e	128	213	340	469	596	725	796	881	1051	1136
−MeOH					564	693	764	849	1019	
Peptide B:	Ac.MeAla.	MeThr.	MeLeu[a].	MeThr.	MeLeu[a].	MeThr.	MeGly.	MeAla.	MeGln.	MeAla.
		OMe		OMe		OMe			NMe_2	
m/e	128	257	384	513	640	769	840	925	1095	1180
−MeOH			352		608	737	808	893	1063	1148

[a] MeLeu or MeIle. *N*-permethylated peptide derivatives having isomeric α-substituents show analogous mass spectra.

for the third residue of each component of this mixture suggested either an *N*-methyl-leucine or its isomer *N*-methylisoleucine (see footnote to table 3). A continuation of this type of stepwise interpretation of the mass spectrum allows the determination of the sequence of the ten amino acid residues from the *N*-terminus. Additional residues were not detected for reasons previously discussed [9]. A second set of *m/e* values in table 3 was produced by loss of methanol from many fragments, a result which is characteristic of peptide derivatives containing *O,N*-dimethylthreonine residues [9].

It should be emphasized that the mass spectrometric sequence determination was accomplished without any knowledge of the chemically determined sequence. The conclusions of each of the two methods are in excellent agreement, and demonstrate that the mass spectrometric technique can reduce significantly the time and effort required for protein-sequence determination. These results also illustrate the ability of mass spectrometry to distinguish the components of a mixture of peptides without their separation; additional mixtures found chemically (table 2) at positions 4 and 6 could not be detected by mass spectrometry because they are present to the extent of less than 5%. The work described here is one of the first examples of the application of mass spectrometry to protein analysis (another case being the sequence determination of an octapeptide from silk fibroin [10]), and the octadecapeptide described above is the largest peptide which has yet been rendered sufficiently volatile for mass spectrometric analysis.

Both methods employed in this work have certain limitations. For mass spectrometry there is the impossibility of distinguishing leucine from isoleucine. The chemical degradation combined with dansylation is not able to distinguish acidic amino acids from their amides. An additional technique, electrophoresis of dansyl-amino acids, is recommended by Gray [3], but electrophoresis is not generally applicable; in our case the dansylpeptides were strongly adsorbed to paper.

The amino acid sequence resulting from the comparative study is homologous to that of human λ-chains, alanine in position 1 of our peptide being homologous with alanine in position 73 of human λ-chains (numbering according to the Atlas of Protein Sequence and Structure [11]). It is expected that the knowledge of the full sequence of the peptide T2bl will contribute to the study of species differences in immunoglobulin polypeptide chains.

References

[1] F.Franěk, B.Keil and F.Šorm, Abstracts of the 5th FEBS Meeting, Prague 1968, p. 173.
[2] D.H.Spackman, W.H.Stein and S.Moore, Anal. Chem. 30 (1958) 1190.
[3] W.R.Gray, in: Methods in Enzymology, vol. 11, ed. C.H.W.Hirs (Academic Press, New York, London, 1967) p. 139.

[4] W.R.Gray, in: Methods in Enzymology, vol. 11, ed. C.H.W.Hirs (Academic Press, New York, London, 1967) p. 469.
[5] J.Novotný and F.Franěk, Chem. Listy 62 (1968) 995.
[6] D.W.Thomas, B.C.Das, S.D.Géro and E.Lederer, Biochem. Biophys. Res. Commun. 32 (1968) 519.
[7] B.C.Das, S.D.Géro and E.Lederer, Biochem. Biophys. Res. Commun. 29 (1967) 211.
[8] J.Coggins and L.Benoiton, Abstracts of Papers, Biol. 18, 156th ACS National Meeting, Atlantic City, 1968.
[9] D.W.Thomas, B.C.Das, S.D.Géro and E.Lederer, Biochem. Biophys. Res. Commun. 32 (1968) 199.
[10] H.R.Morris, A.J.Geddes and G.N.Graham, Paper presented at the 489th Meeting of the Biochemical Society, London, 1968 (Biochem. J., 1969, in press).
[11] M.O.Dayhoff, R.V.Eck, Atlas of Protein Sequence and Structure 1967–1968 (National Biomedical Research Foundation, Silver Springs, 1968).